OIPE IAP13
SEP 09 2005
PATENT & TRADEMARK OFFICE

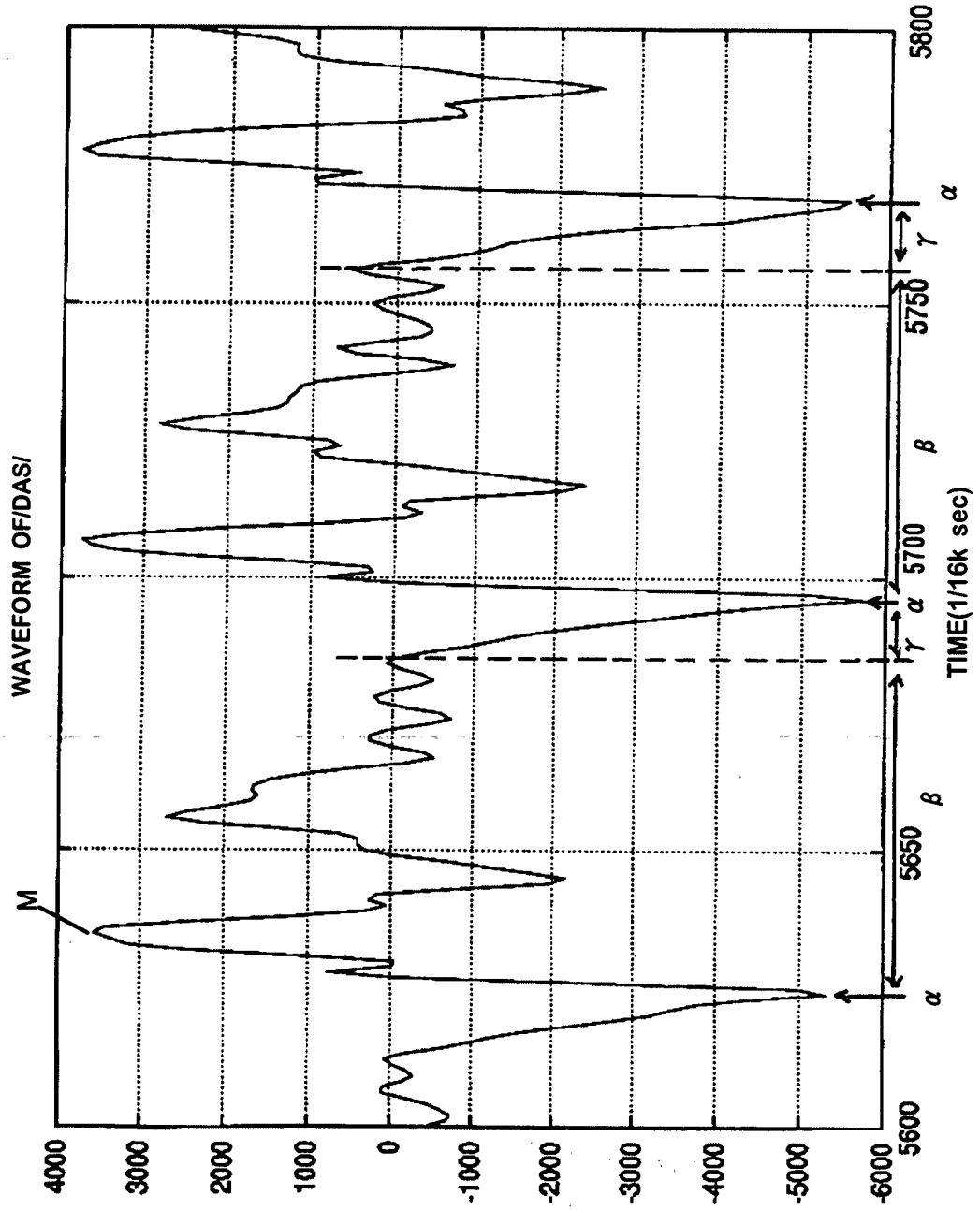


FIG.1

FIG.2

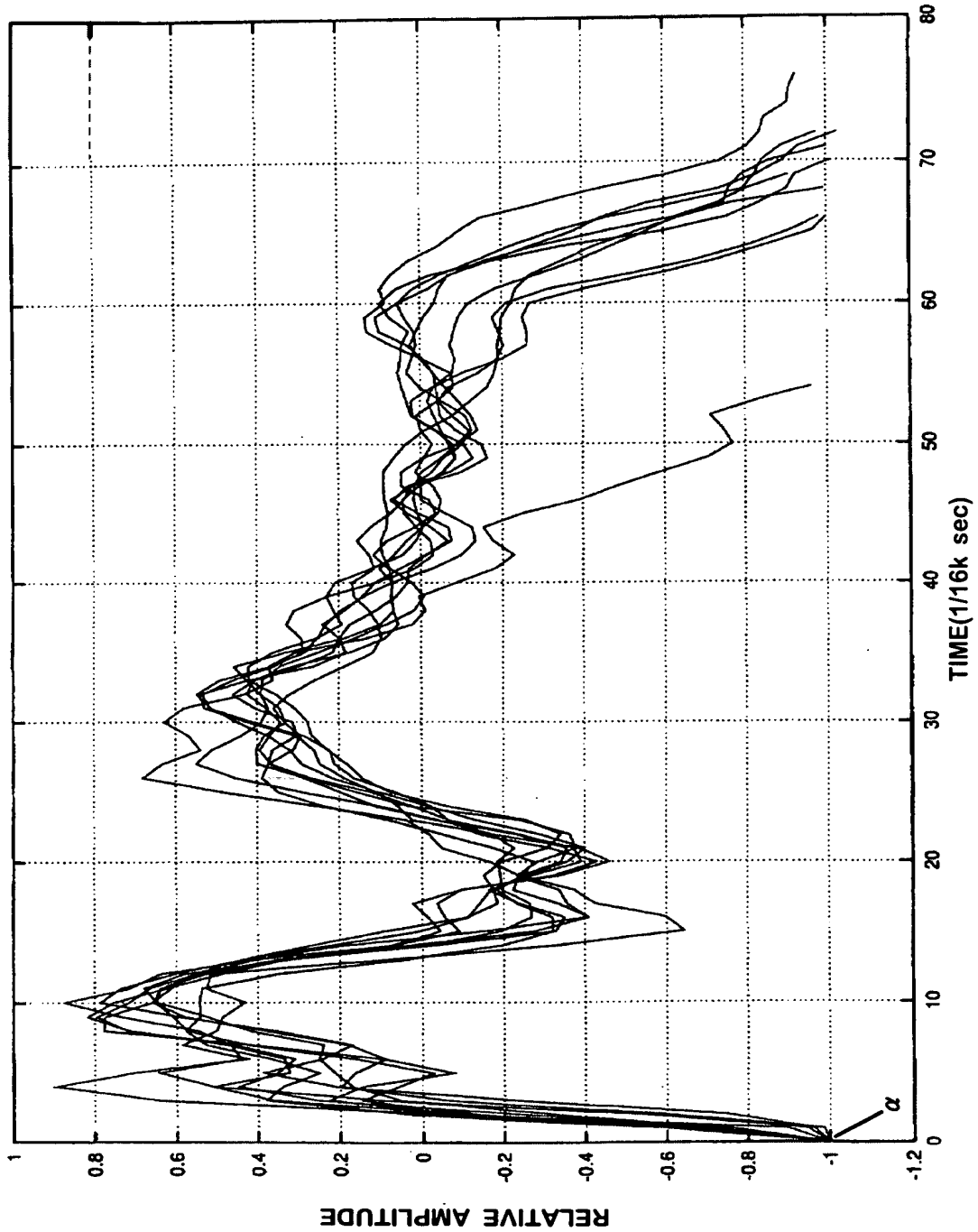

RELATIVE AMPLITUDE
TIME(1/16k sec)
α
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
-1.2
0
10
20
30
40
50
60
70
80

FIG.3

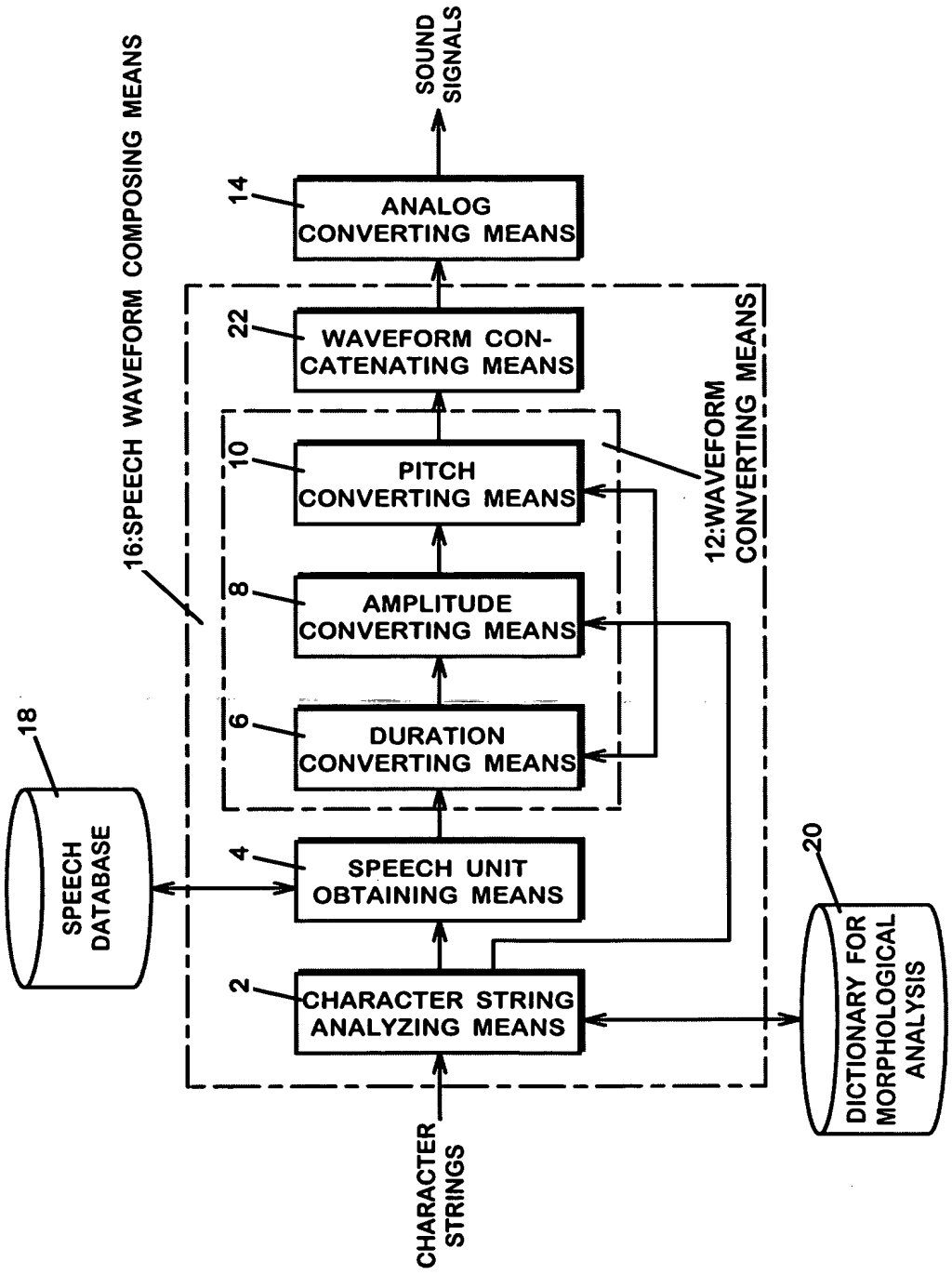

SOUND SIGNALS
14
ANALOG CONVERTING MEANS
22
WAVEFORM CONCATENATING MEANS
10
PITCH CONVERTING MEANS
8
AMPLITUDE CONVERTING MEANS
6
DURATION CONVERTING MEANS
4
SPEECH UNIT OBTAINING MEANS
2
CHARACTER STRING ANALYZING MEANS
16:SPEECH WAVEFORM COMPOSING MEANS
12:WAVEFORM CONVERTING MEANS
18
SPEECH DATABASE
20
DICTIONARY FOR MORPHOLOGICAL ANALYSIS
CHARACTER STRINGS

FIG.4

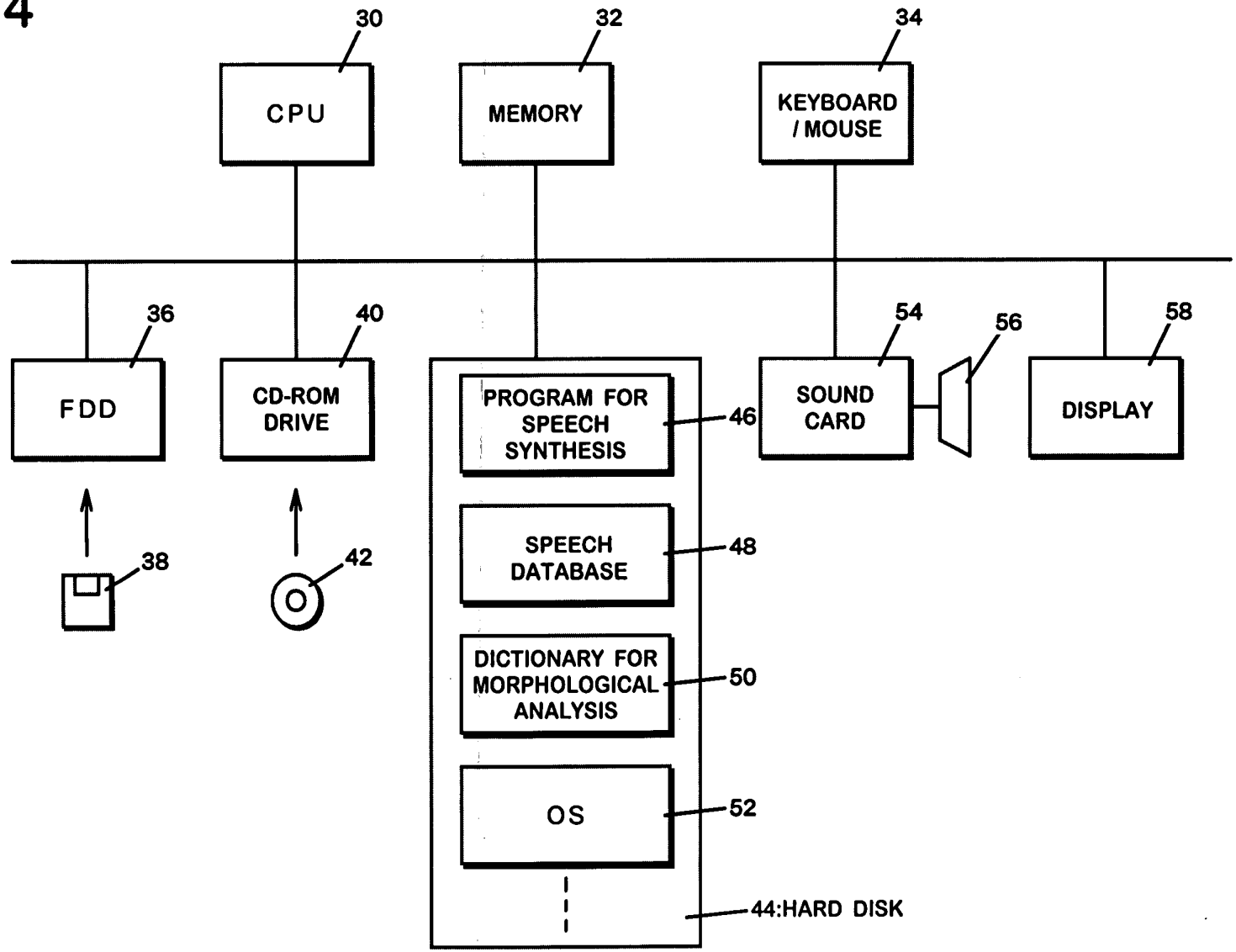

30
32
34
CPU
MEMORY
KEYBOARD / MOUSE
36
40
54
56
58
FDD
CD-ROM DRIVE
SOUND CARD
DISPLAY
PROGRAM FOR SPEECH SYNTHESIS
46
SPEECH DATABASE
48
DICTIONARY FOR MORPHOLOGICAL ANALYSIS
50
OS
52
38
42
44:HARD DISK

FIG.5

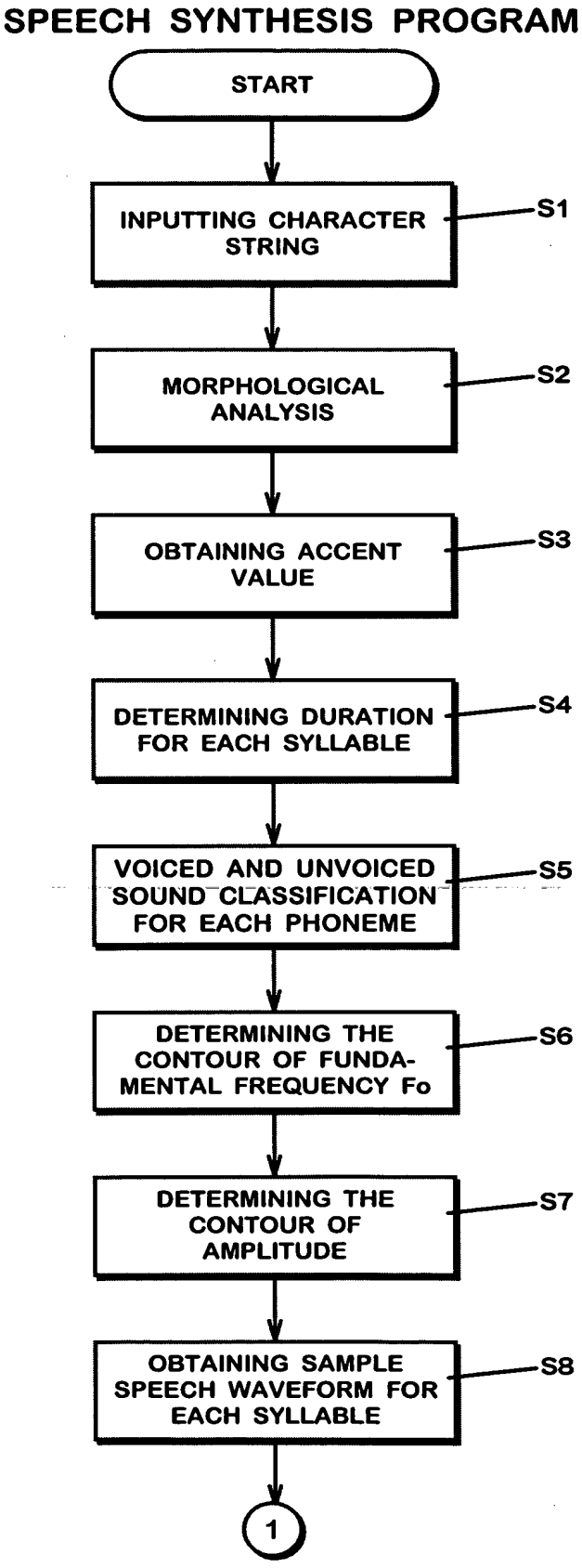

SPEECH SYNTHESIS PROGRAM
START
INPUTTING CHARACTER STRING
S1
MORPHOLOGICAL ANALYSIS
S2
OBTAINING ACCENT VALUE
S3
DETERMINING DURATION FOR EACH SYLLABLE
S4
VOICED AND UNVOICED SOUND CLASSIFICATION FOR EACH PHONEME
S5
DETERMINING THE CONTOUR OF FUNDA-MENTAL FREQUENCY Fo
S6
DETERMINING THE CONTOUR OF AMPLITUDE
S7
OBTAINING SAMPLE SPEECH WAVEFORM FOR EACH SYLLABLE
S8
1

FIG.6

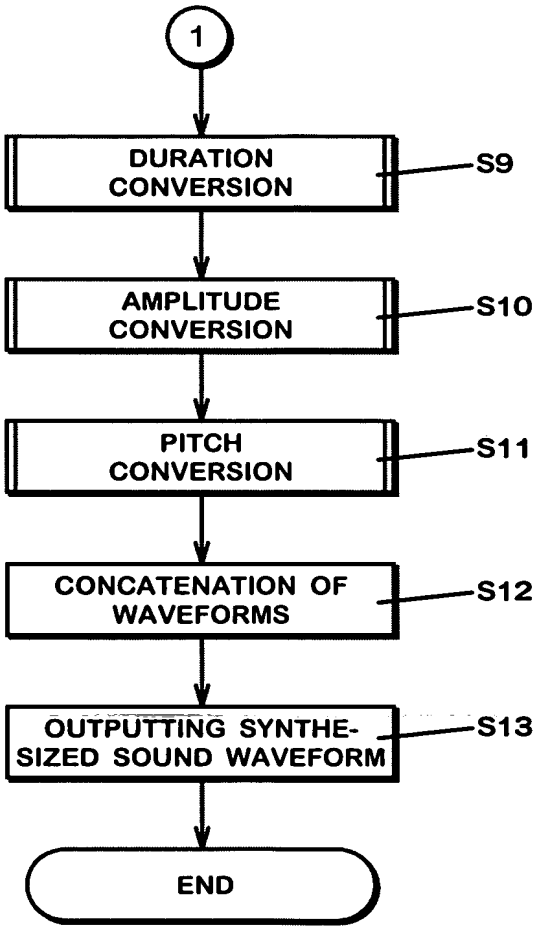

1
DURATION CONVERSION
S9
AMPLITUDE CONVERSION
S10
PITCH CONVERSION
S11
CONCATENATION OF WAVEFORMS
S12
OUTPUTTING SYNTHE-SIZED SOUND WAVEFORM
S13
END

FIG.7

PROGRAM FOR PITCH CONVERSION PROCESS

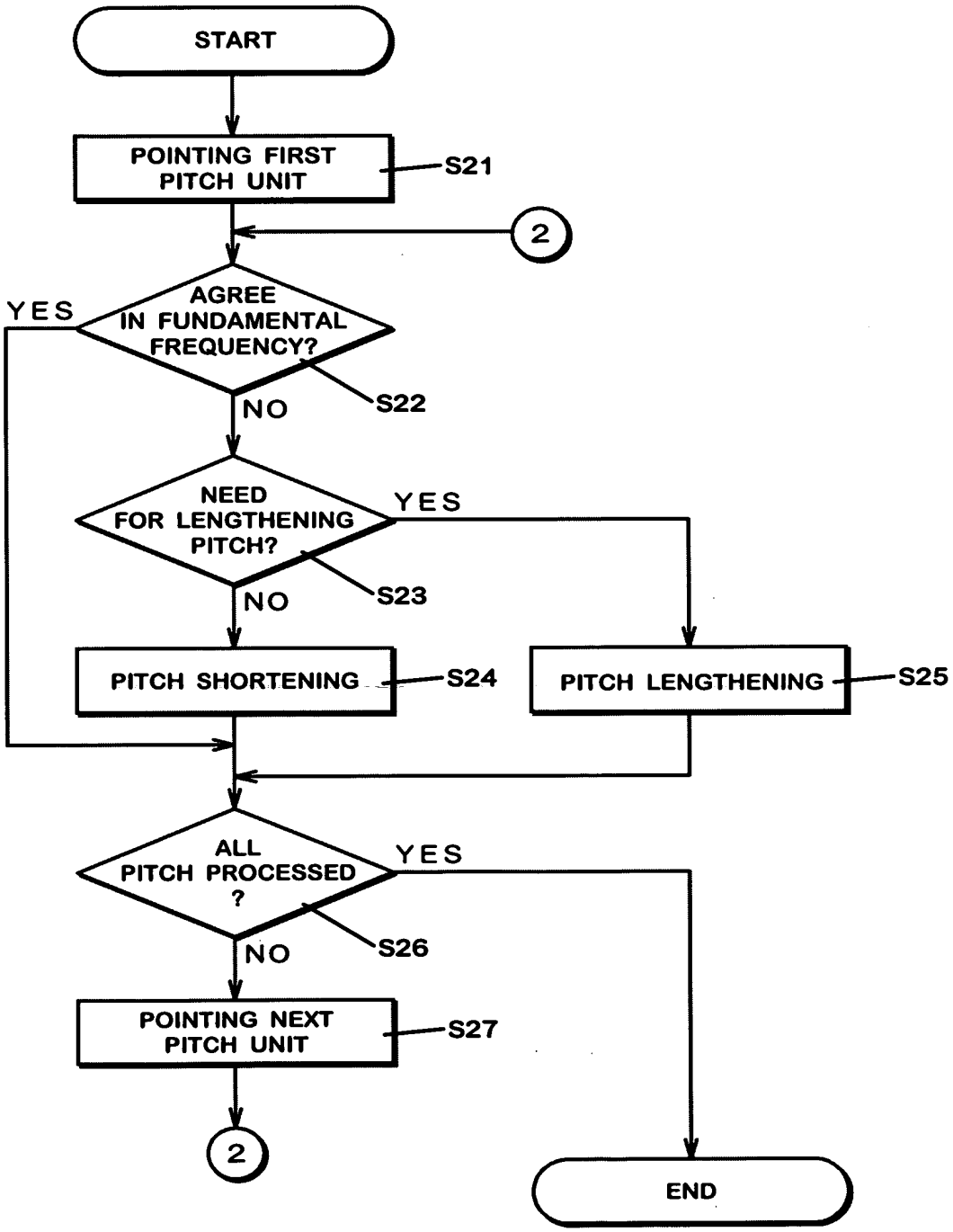


FIG.8

WORD DICTIONARY

WORD	PART OF SPEECH	READING	----	SYLLABLE	ACCENT VALUE
桜	NOUN	s a k u r a	----	s a	1
				k u	5
				r a	0
さく	VERB	s a k u	----	s a	5
				k u	0
⋮	⋮	⋮		⋮	⋮
が	POST POSITIONAL PARTICAL OF JAPANESE	g a	----	g a	2
⋮	⋮	⋮		⋮	⋮

FIG.9

DICTIONARY OF SYLLABLE DURATION

SYLLABLE	DURATION [ms]
a	110
i	114
u	90
⋮	⋮
ko	188
⋮	⋮

FIG.10

SYLLABLE	CHAIN OF SYLLABLE SEQUENCE	[ms] TIME	[ms] SYLLABLE DURATION	ACCENT VALUE
k o	k o N	0	188	0
N	N n i	188	92	5
n i	n i c h	280	212	5
c h i	c h i w	492	178	0
w a	w a	670	166	7

FIG.11

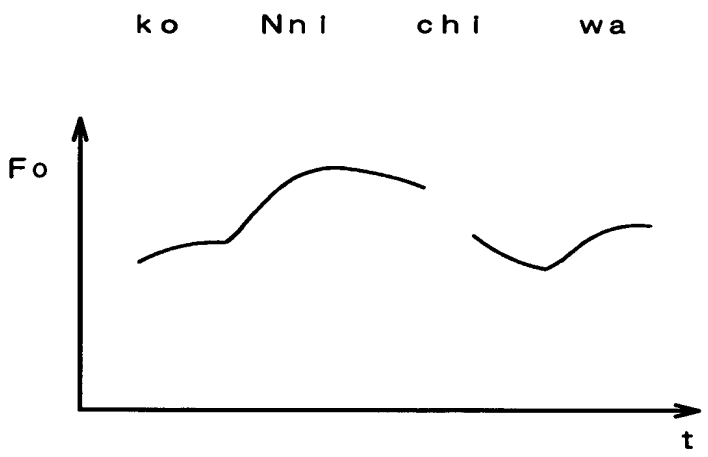


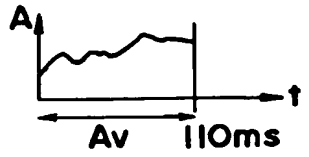
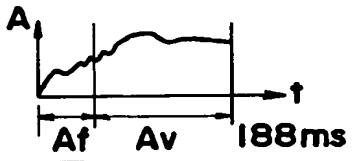
FIG.12

DICTIONARY OF VOICED / UNVOICED SOUNDS
FOR CONSONANTS / VOWELS

PHONEME	INDEX
a	V
i	V
⋮	⋮
k	CU
⋮	⋮
b	CV
⋮	⋮

FIG.13

DICTIONARY OF SOUND SOURCE AMPLITUDE

SYLLABLE	SOUND SOURCE AMPLITUDE
a	
⋮	⋮
ko	
⋮	⋮

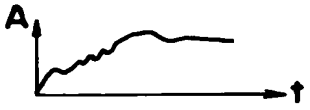
Av : VOICED SOUND SOURCE AMPLITUDE
Af : UNVOICED SOUND SOURCE AMPLITUDE

FIG.14

BEST AVAILABLE COPY

Ko N Do No Ni Chi

CONTOUR OF
SOUND SOURCE AMPLITUDE OF / ko /



CONTOUR OF
FUNDAMENTAL FREQUENCY OF / ko /

DURATION OF / ko / 123ms

PITCH MARK 0、15、30、…

ZERO CROSSING MARK 12、27、42、…

FIG.15

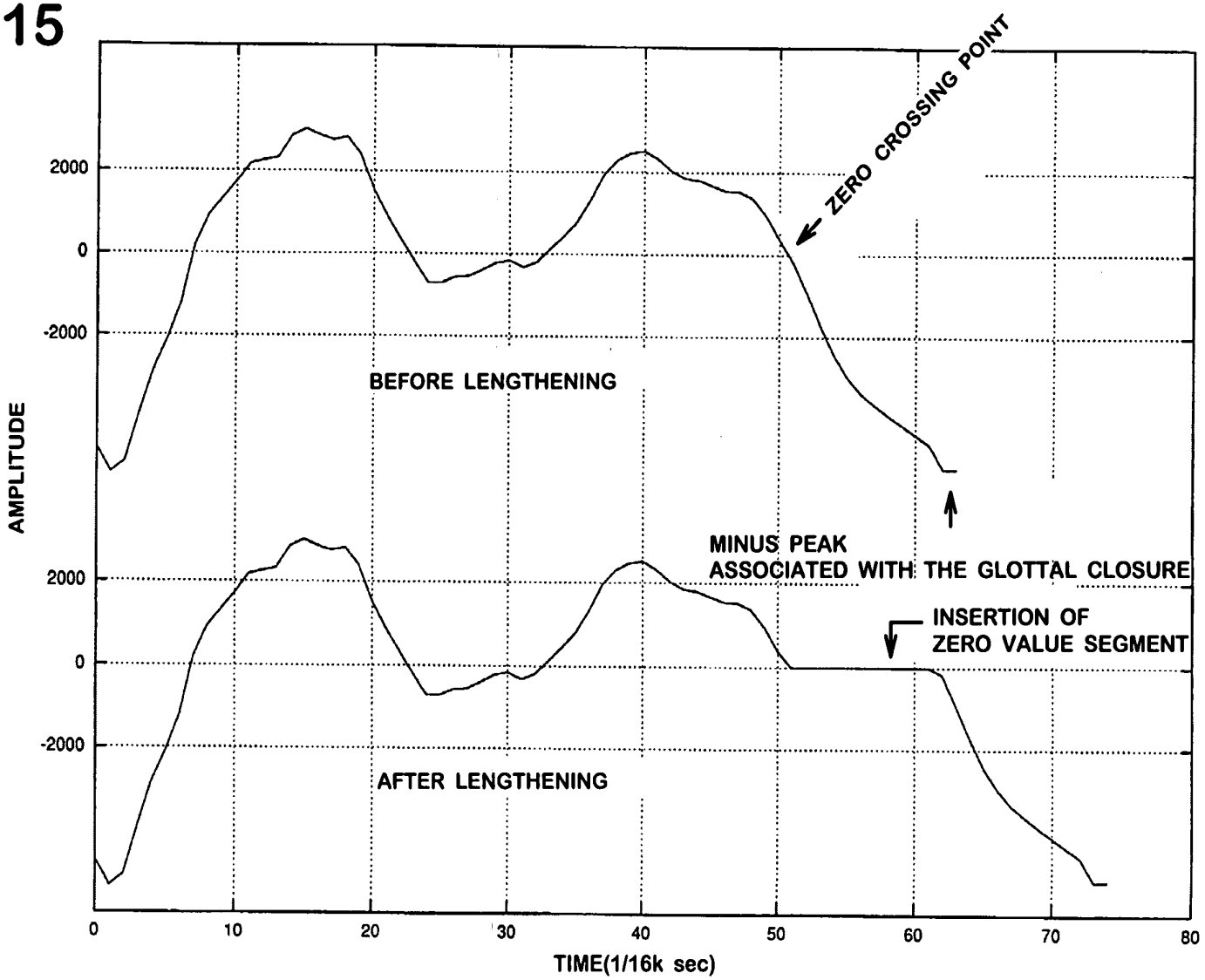

ZERO CROSSING POINT
2000
0
-2000
BEFORE LENGTHENING
AMPLITUDE
MINUS PEAK
ASSOCIATED WITH THE GLOTTAL CLOSURE
INSERTION OF
ZERO VALUE SEGMENT
2000
0
-2000
AFTER LENGTHENING
0
10
20
30
40
50
60
70
80
TIME(1/16k sec)

FIG.16

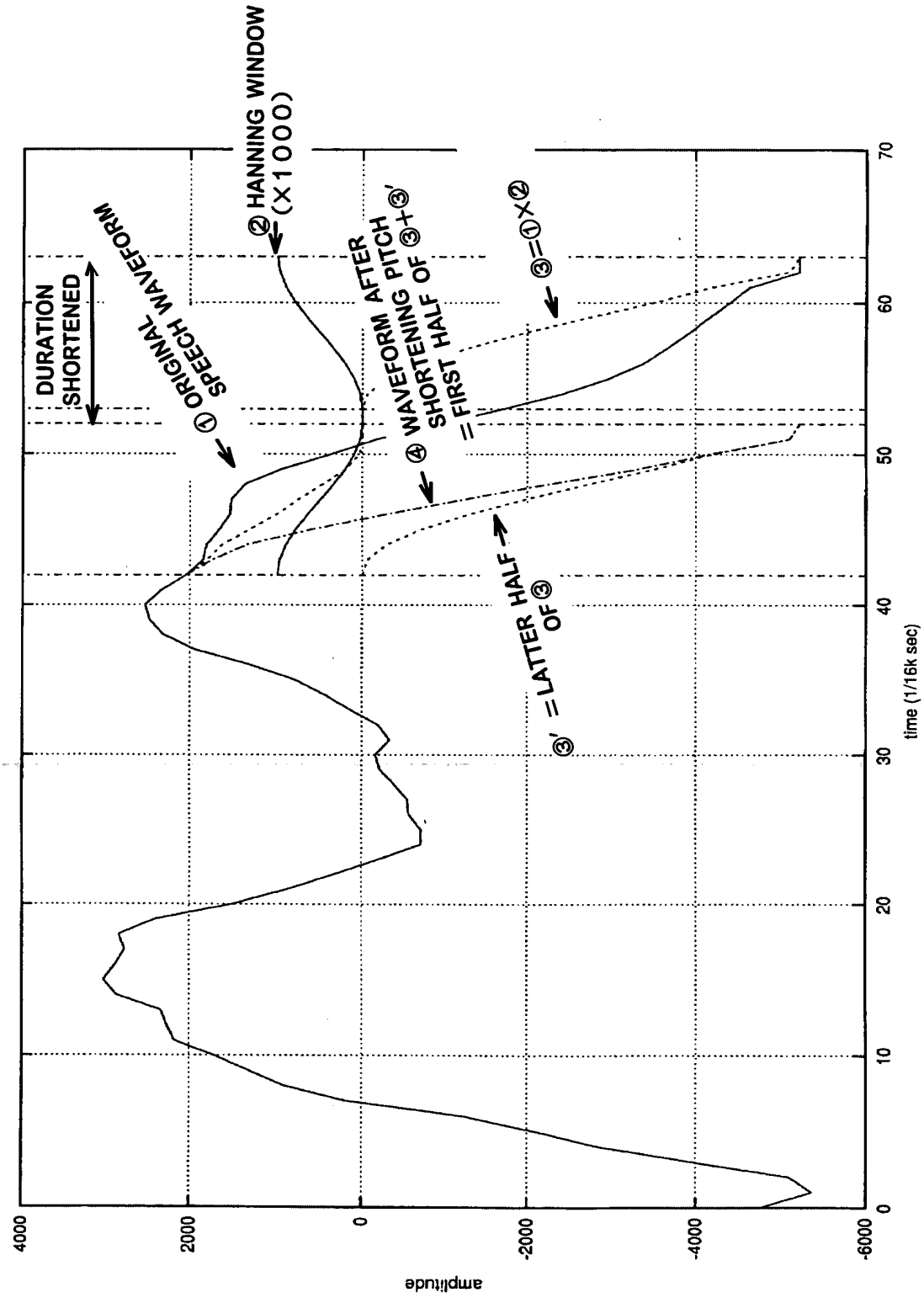

DURATION SHORTENED
① ORIGINAL WAVEFORM SPEECH
② HANNING WINDOW (×1000)
③=①×②
④ WAVEFORM AFTER SHORTENING PITCH OF ③+③′ = FIRST HALF
③′ = LATTER HALF OF ③
amplitude
time (1/16k sec)
4000
2000
0
-2000
-4000
-6000
0
10
20
30
40
50
60
70

FIG.17

SYLLABLE WEIGHT	TYPE OF SYLLABLE WEIGHT	SYLLABLE STRUCTURE	EXAMPLE
1	LIGHT SYLLABLE	(C) (y) V	ka, sa, ta, na, ha, ma, ya, ra,... a, i, u, e, o, ... che, pya,....
2	HEAVY SYLLABLE	(C) (y) V R (C) (y) V J (C) (y) V N (C) (y) V Q	to: , ya:, kyu:, pyu:,... kai, nou, ai, ui, pyua,... kaN, aN, myaN, chaN,... chuQ, ryaQ, jaQ, hyaQ,...
3 OR MORE	SUPER HEAVY SYLLABLE	(C) (y) V R N (C) (y) V R Q (C) (y) V J N (C) (y) V J Q (C) (y) V N Q AND THE LIKE	che:N, ju:N, a:N, ... u:Q, che:Q,... saiN, pauN, auN, chaiN,... kaiQ, daiQ, kyaiQ, uiQ,... doNQ, uNQ, chaNQ,...

C : CONSONANT (EXCLUDING Q, Y AND N)
y : SEMI VOWEL
V : VOWEL (EXCLUDING R AND J)
R : LONG VOWEL
J : THE SECOND ELEMENT OF DIPHTHONG
Q : GEMINATED SOUND (JAPANESE SOKUON)
N : SYLLABIC NASAL
(X) : SYLLABLE WEIGHT IS INDEPENDENT OF X

FIG.18

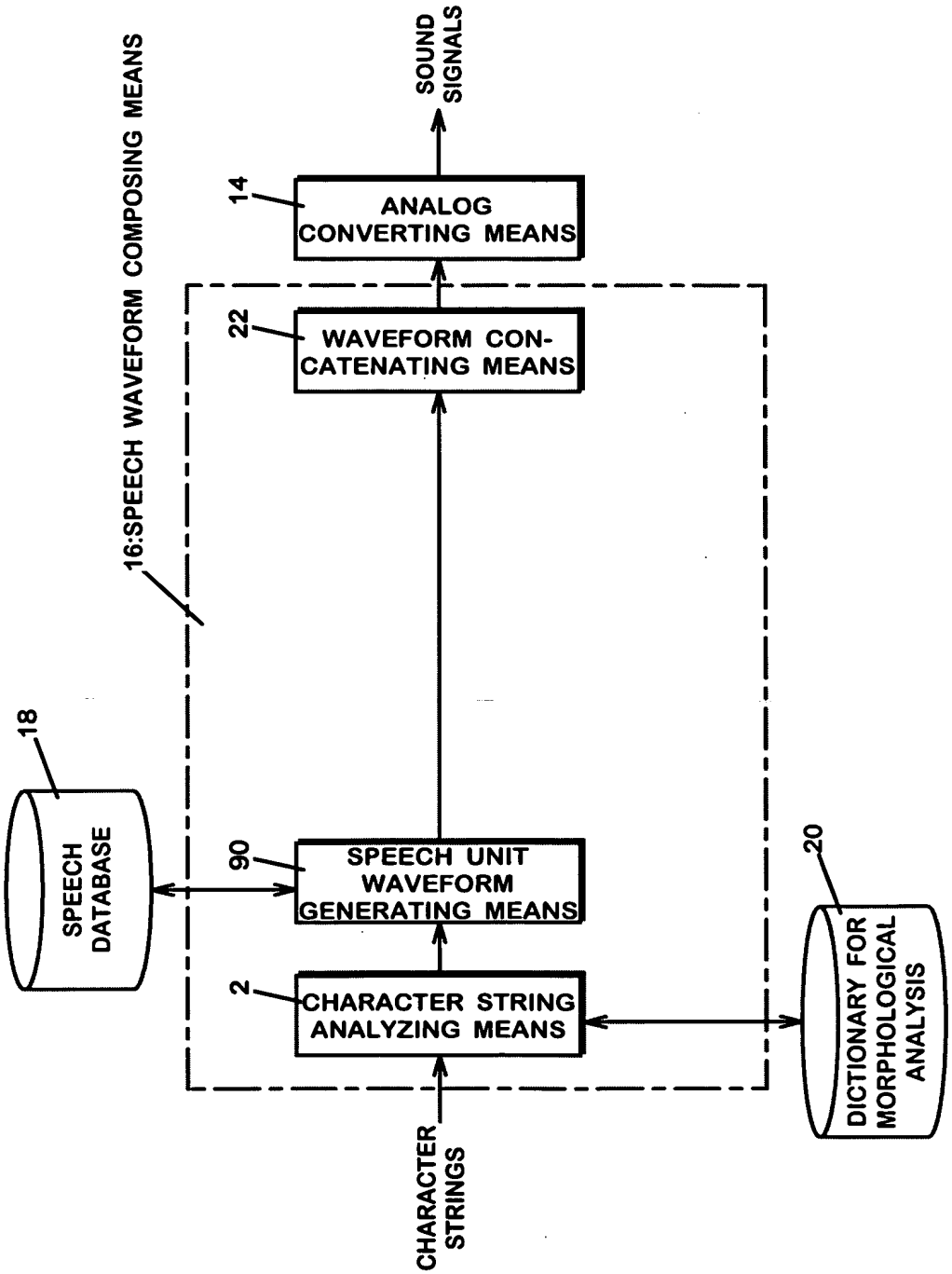

SOUND SIGNALS
14
ANALOG CONVERTING MEANS
22
WAVEFORM CONCATENATING MEANS
16:SPEECH WAVEFORM COMPOSING MEANS
18
SPEECH DATABASE
90
SPEECH UNIT WAVEFORM GENERATING MEANS
2
CHARACTER STRING ANALYZING MEANS
20
DICTIONARY FOR MORPHOLOGICAL ANALYSIS
CHARACTER STRINGS

FIG.19

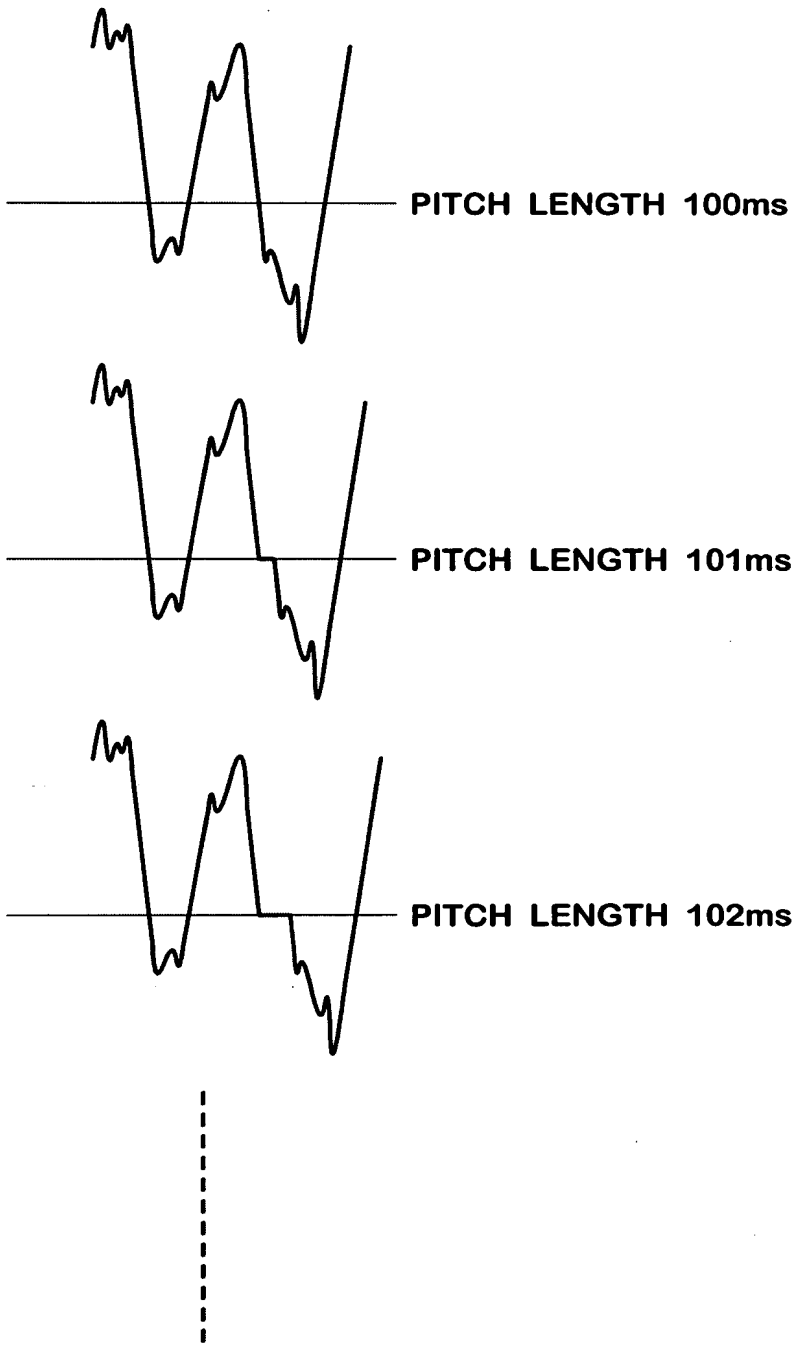


FIG.20

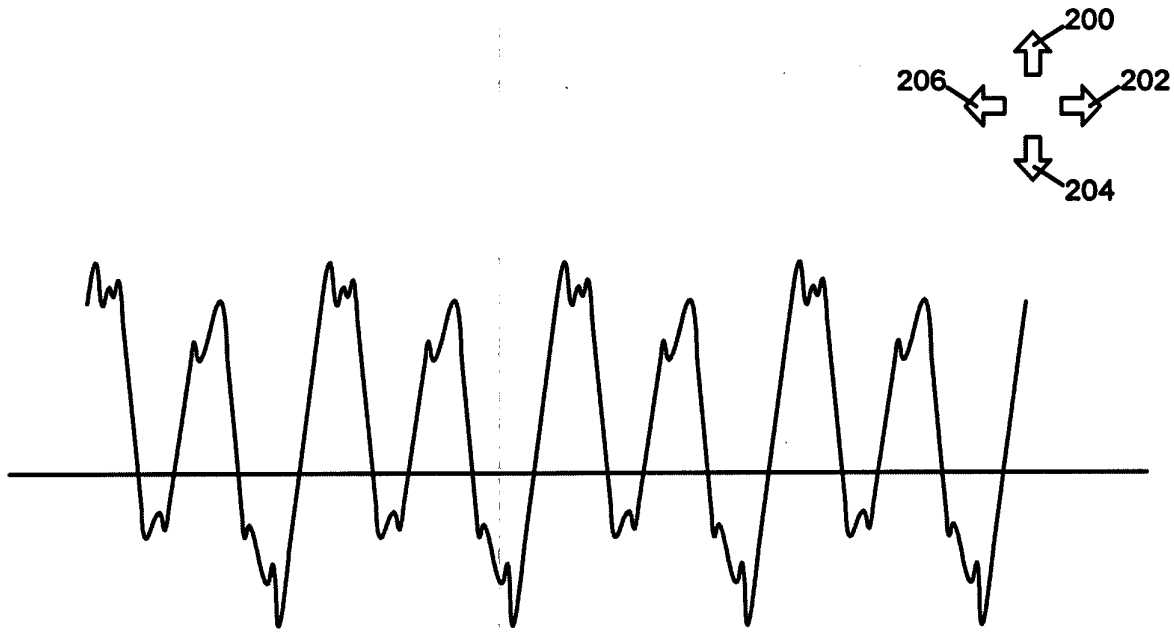


FIG.21

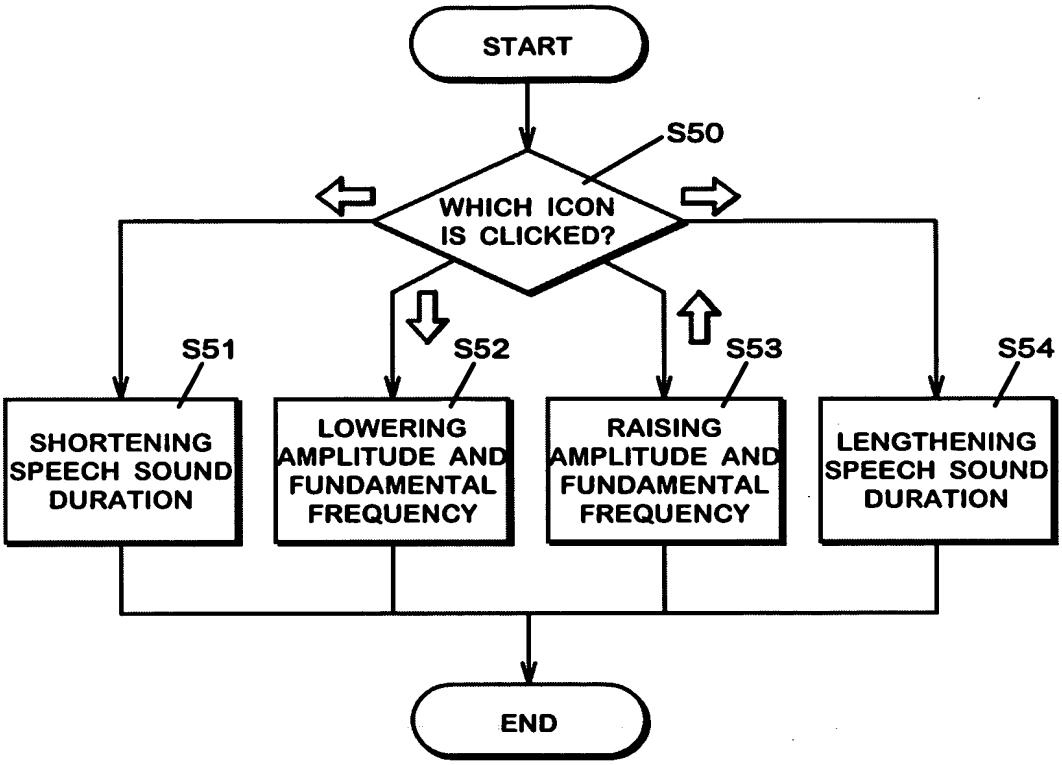

START
S50
WHICH ICON IS CLICKED?
S51
S52
S53
S54
SHORTENING SPEECH SOUND DURATION
LOWERING AMPLITUDE AND FUNDAMENTAL FREQUENCY
RAISING AMPLITUDE AND FUNDAMENTAL FREQUENCY
LENGTHENING SPEECH SOUND DURATION
END

FIG.22

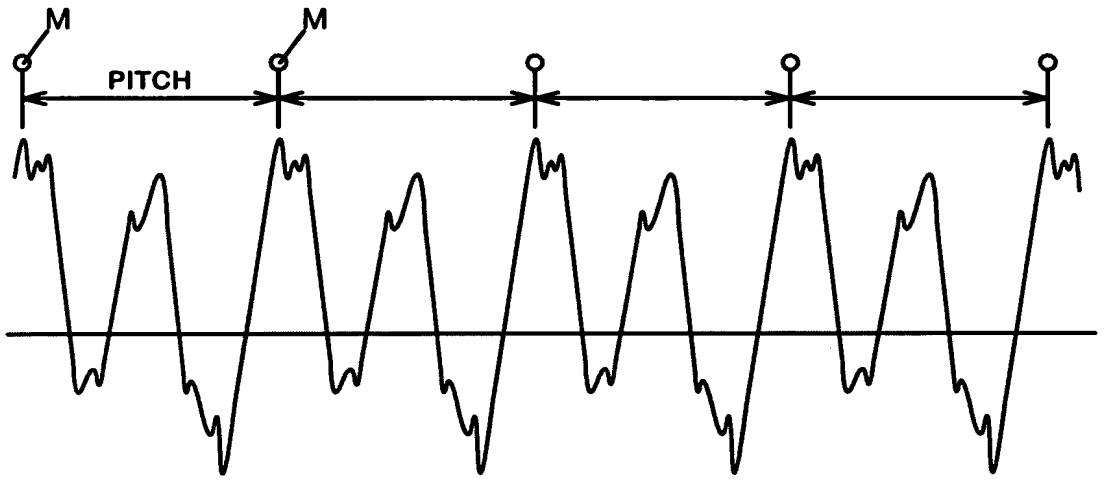


FIG.23

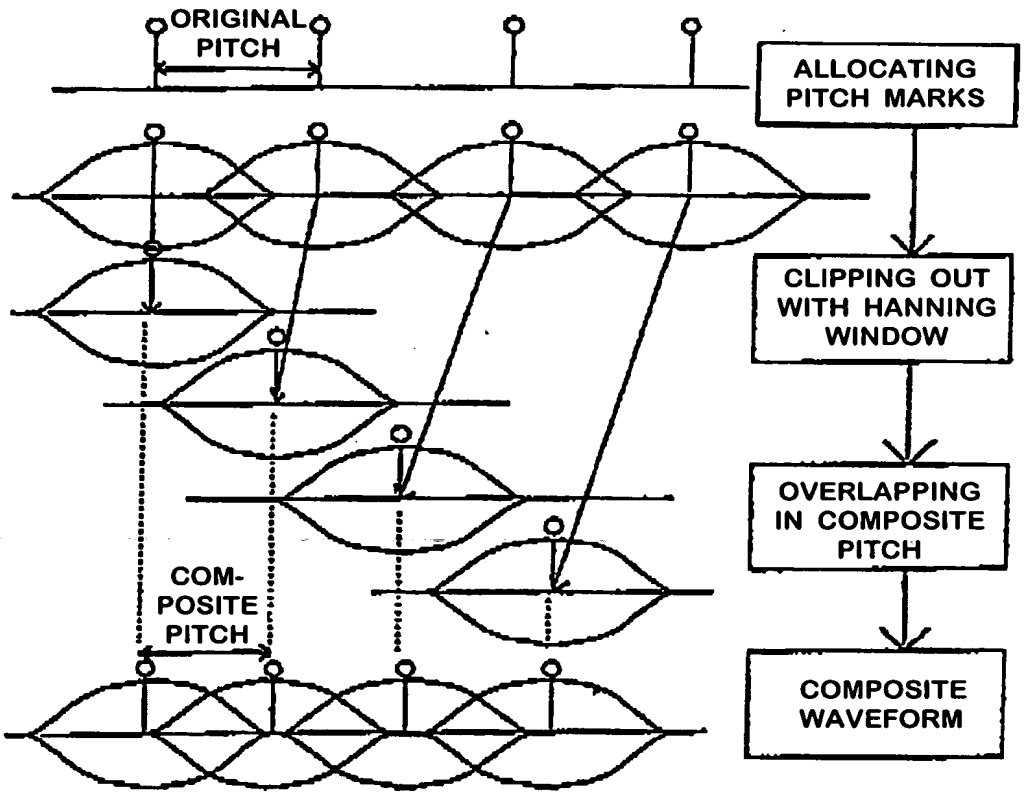

ORIGINAL PITCH
ALLOCATING PITCH MARKS
CLIPPING OUT WITH HANNING WINDOW
OVERLAPPING IN COMPOSITE PITCH
COMPOSITE WAVEFORM
COM-POSITE PITCH

FIG.24

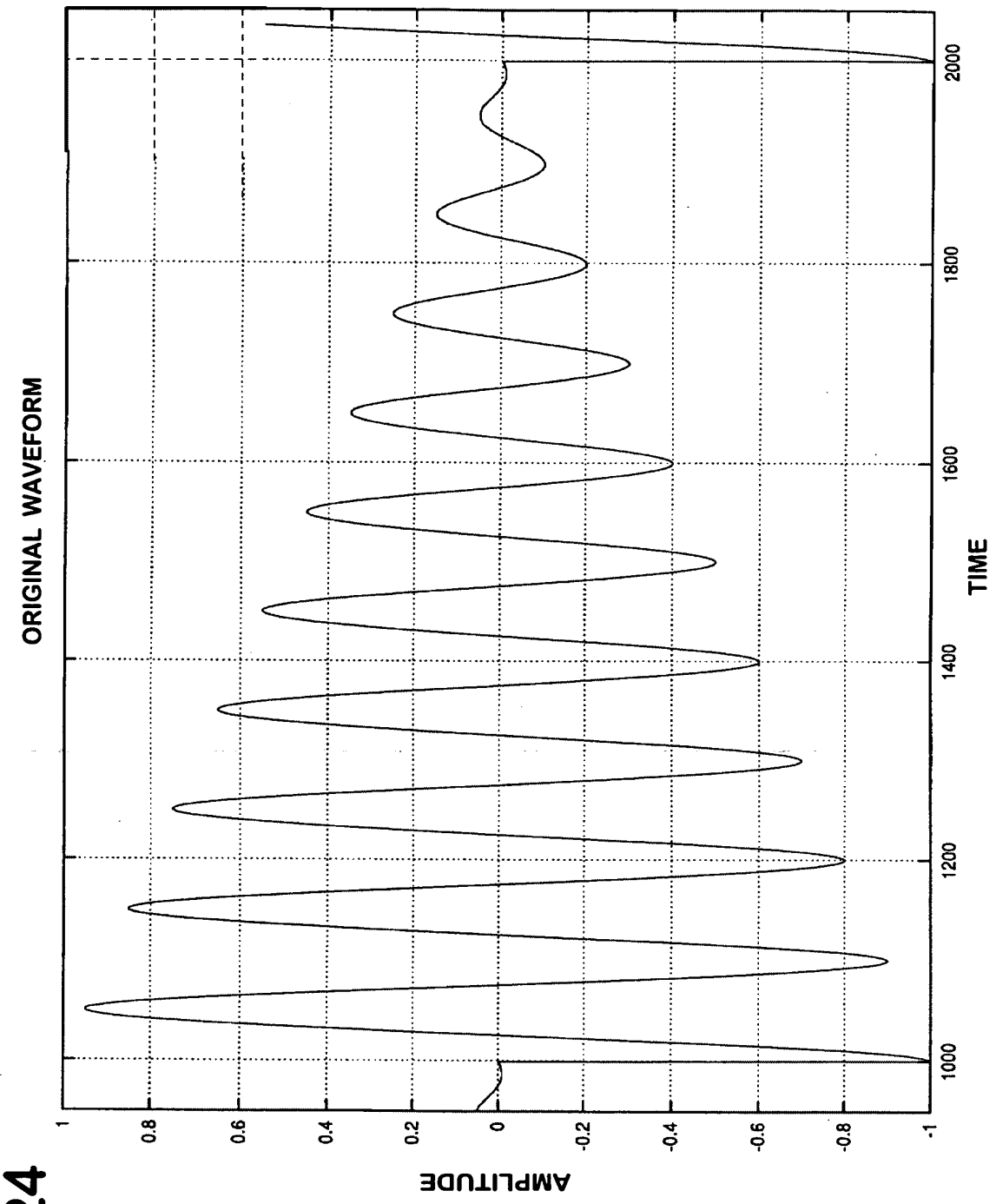

ORIGINAL WAVEFORM
AMPLITUDE
TIME
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
1000
1200
1400
1600
1800
2000

FIG.25

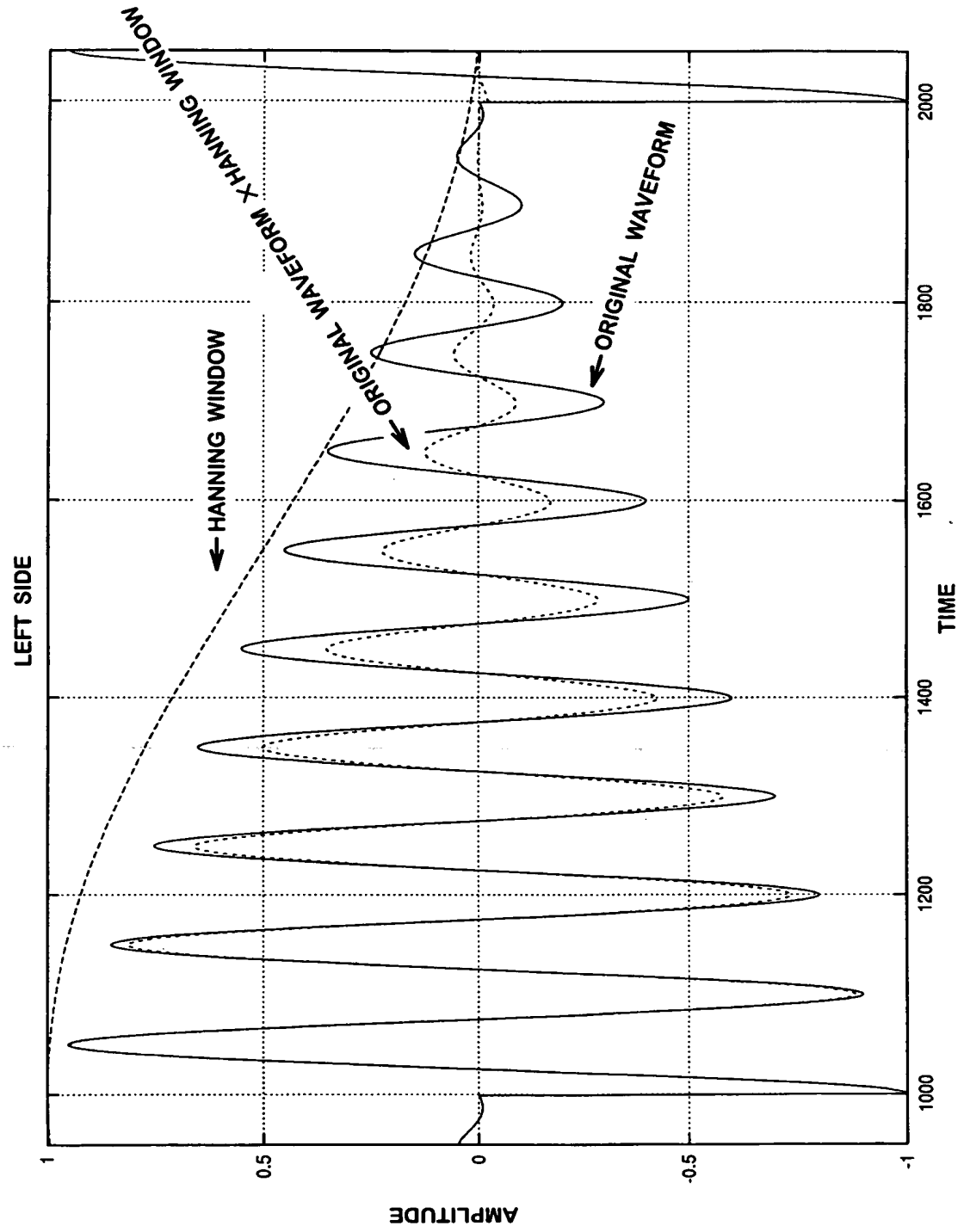

LEFT SIDE
HANNING WINDOW
ORIGINAL WAVEFORM X HANNING WINDOW
ORIGINAL WAVEFORM
AMPLITUDE
TIME
1
0.5
0
-0.5
-1
1000
1200
1400
1600
1800
2000

FIG.26

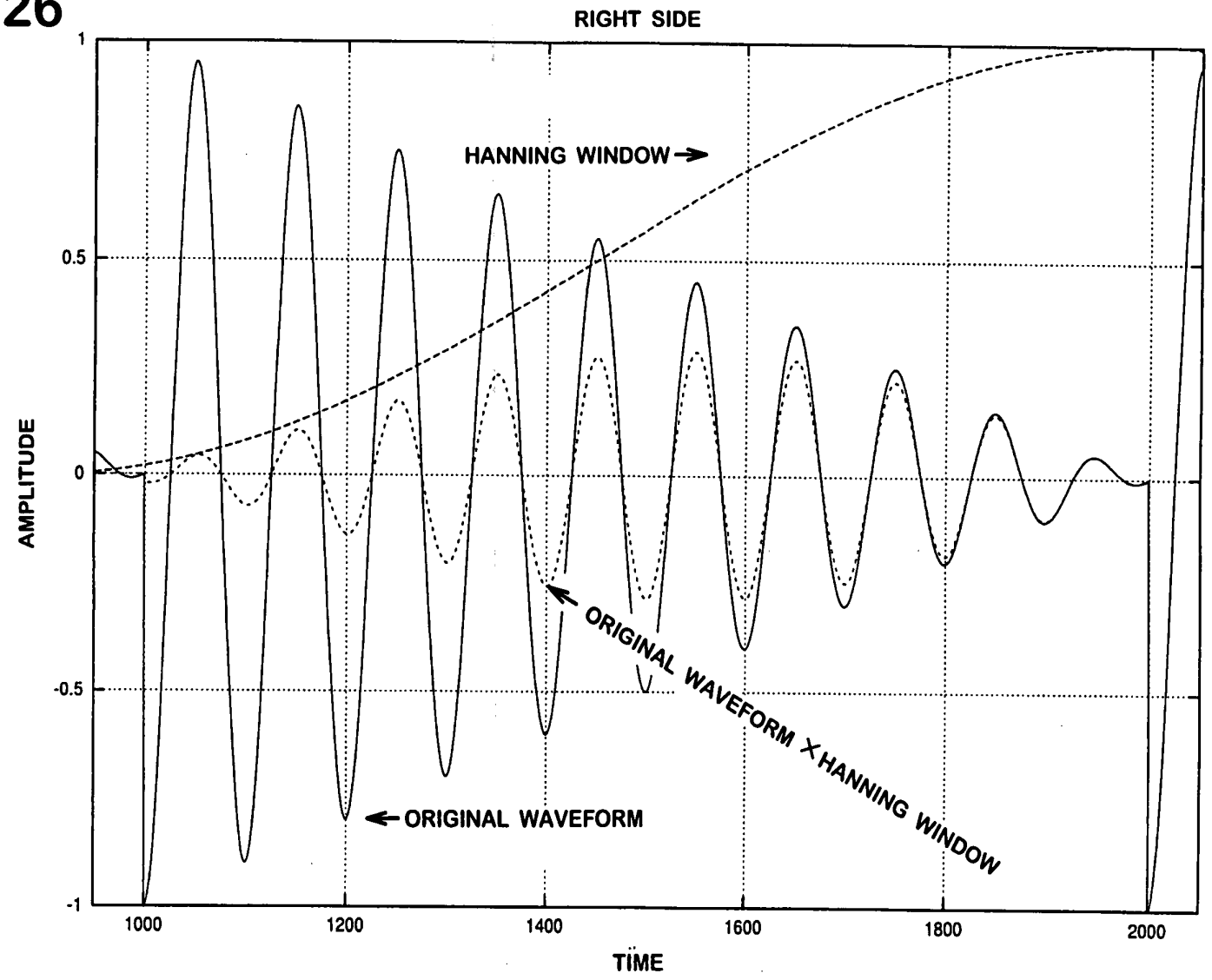

RIGHT SIDE
HANNING WINDOW
ORIGINAL WAVEFORM X HANNING WINDOW
ORIGINAL WAVEFORM
AMPLITUDE
TIME
1
0.5
0
-0.5
-1
1000
1200
1400
1600
1800
2000

FIG.27

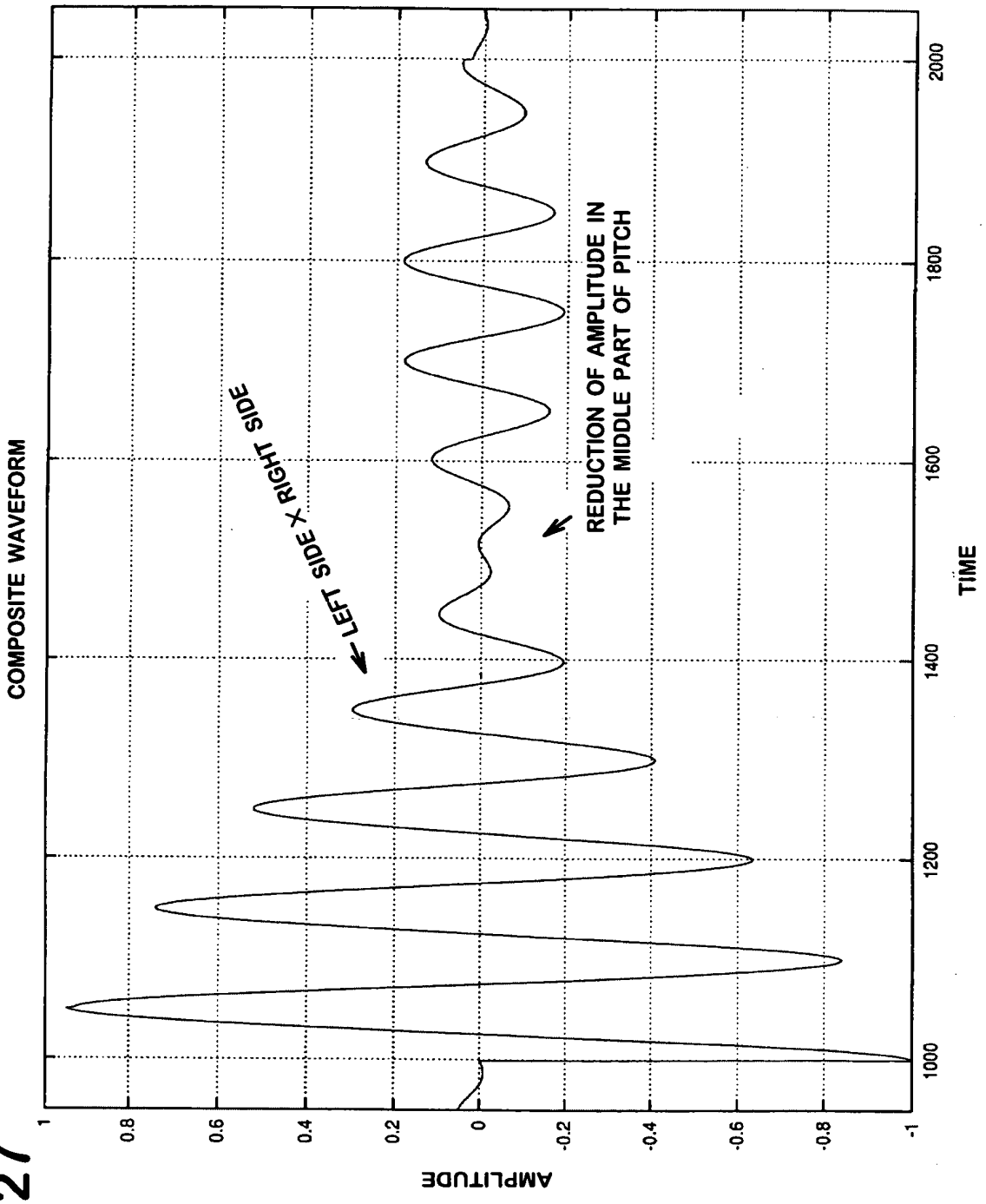
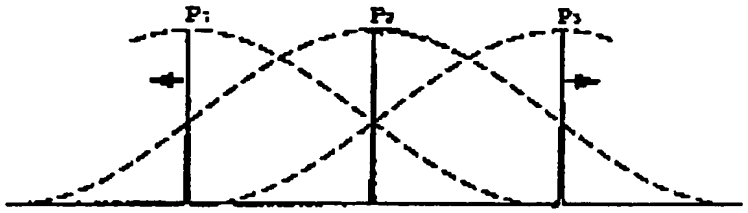
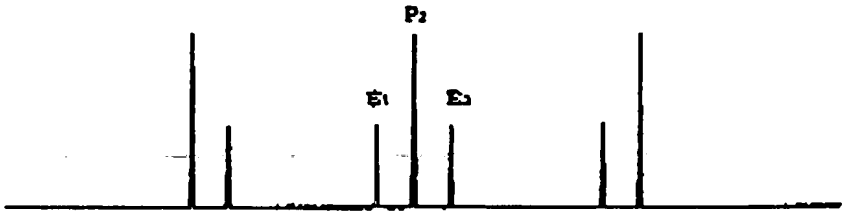

COMPOSITE WAVEFORM
LEFT SIDE X RIGHT SIDE
REDUCTION OF AMPLITUDE IN THE MIDDLE PART OF PITCH
AMPLITUDE
TIME
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
1000
1200
1400
1600
1800
2000

FIG.28

ORIGINAL SIGNALS (SOLID LINE) AND WINDOW FUNCTION (BROKEN LINE)



SIGNALS AFTER CONVERSION OF Fo